JOURNAL OF THE AMERICAN WATER RESOURCES ASSOCIATION
VOL. 35, NO. 2 AMERICAN WATER RESOURCES ASSOCIATION APRIL 1999

ACCEPTING THE STANDARDIZED PRECIPITATION INDEX: A CALCULATION ALGORITHM[1]

Nathaniel B. Guttman[2]

ABSTRACT: The Palmer Drought Severity Index (PDSI) has been calculated for about 30 years as a means of providing a single measure of meteorological drought severity. It was intended to retrospectively look at wet and dry conditions using water balance techniques. The Standardized Precipitation Index (SPI) is a probability index that was developed to give a better representation of abnormal wetness and dryness than the Palmer indices. Before the user community will accept the SPI as an alternative to the Palmer indices, a standard method must be developed for computing the index. Standardization is necessary so that all users of the index will have a common basis for both spatial and temporal comparison of index values. If different probability distributions and models are used to describe an observed series of precipitation, then different SPI values may be obtained. This article describes the effect on the SPI values computed from different probability models as well as the effects on dry event characteristics. It is concluded that the Pearson Type III distribution is the "best" universal model, and that the reliability of the SPI is sample size dependent. It is also concluded that because of data limitations, SPIs with time scales longer than 24 months may be unreliable. An internet link is provided that will allow users to access Fortran 77 source code for calculating the SPI.
(KEY TERMS: Palmer Index; Standardized Precipitation Index; drought; precipitation probability distributions.)

INTRODUCTION

The Palmer Drought Severity Index (PDSI) has been calculated for about 30 years as a means of providing a single measure of meteorological drought severity. It was intended to retrospectively look at wet and dry conditions from a water balance viewpoint (Palmer, 1965). Variations of the index include the modified PDSI (PMDI), which is designed for real-time use; the Palmer Hydrologic Drought Index, which is used for water supply monitoring; and the Z Index, which is a measure of an individual month's wetness or dryness. These Palmer indices are based on a monthly water balance accounting scheme that involves precipitation, evapotranspiration, runoff and soil moisture. They have been used extensively for monitoring drought and for making operational water management decisions.

The Standardized Precipitation Index (SPI) was developed by McKee *et al.* (1993; 1995) to give a better representation of abnormal wetness and dryness than the Palmer indices. The SPI is probability based and was designed to be a spatially invariant indicator of drought that recognizes the importance of time scales in the analysis of water availability and water use. It is essentially a standardizing transform of the probability of the observed precipitation. It can be computed for a precipitation total observed over any duration desired by a user. Short term durations on the order of months (or even weeks) may be important to agricultural interests while very long term durations spanning years may be important to water supply management interests.

Guttman (1997) compared spectral characteristics of the PMDI, which is used for real-time monitoring, to those of the SPI for time scales ranging from one month to three years. He recommended that the SPI be used as the primary drought index because it is simple, spatially invariant in its interpretation, and probabilistic so that it can be used in risk and decision analysis. It also can be tailored to time periods of concern to a user. He noted that in contrast, the modified PDSI is very complex, spatially variant, difficult to interpret, and has an inherent fixed time scale of about 9-12 months. Spatial and temporal comparisons made by users of the Palmer values may therefore be misleading and lead to erroneous conclusions. The

[1]Paper No. 97156 of the *Journal of the American Water Resources Association*. **Discussions are open until December 1, 1999.**
[2]National Climatic Data Center, 151 Patton Ave., Asheville, North Carolina 28801-5001 (E-Mail: nguttman@ncdc.noaa.gov).

reader is referred to Alley (1984), Heddinghaus and Sabol (1991), Karl (1983; 1986), Guttman (1991), and Guttman *et al.* (1992) for discussions of the characteristics, advantages and disadvantages of the Palmer indices.

Prior to acceptance of the SPI by the user community, a standard method must be used for computing the index. If different probability distributions and models are used to describe an observed series of precipitation data, then different SPI values may be obtained. This article describes the changes in the SPI values when computed from different probability models and specifically addresses dry event characteristics. Recommendations are also given for standardizing the SPI calculations.

THE SPI

The first step in calculating the SPI is to determine a probability density function that describes the long term time series of precipitation observations. The series can be for any time duration (i.e., running series of total precipitation for 1 month, 2 months, 6 months, 1 year, 3 years, etc.) Once the probability density function is determined, the cumulative probability of an observed precipitation amount is computed. The inverse normal (Gaussian) function, with mean zero and variance one, is then applied to the cumulative probability. The result is the SPI.

For a given time scale, SPI values are positive (negative) for greater (less) than median precipitation. A value of zero corresponds to the median precipitation. The magnitude of the departure from zero is a probabilistic measure of the severity of a wet or dry event that can be used for risk assessment. The time series of the SPI can be used for drought monitoring by setting application-specific thresholds of the SPI for defining drought beginning and ending times. Accumulated values of the SPI can be used to analyze drought severity. It should be noted that while this article concentrates primarily on drought, the SPI is equally effective as a measure of both wetness and aridity.

Theoretically, the SPI is unbounded. Practically, however, the number of observations of the precipitation data, which is generally less than 100 for a given month, season or other time period, for locations in the United States, suggests bounds of ± 3.09. These bounds correspond to probabilities of .999 and .001 and return periods of one in 1,000. Estimation of more extreme probabilities, based on sample sizes of only 100, is likely to be inaccurate. Some people may also argue that bounds of ± 2.33, which correspond to probabilities of .99 and .01 and return periods of one in 100 are better. In either case, however, SPI values higher (lower) than 2.00 (-2.00) can be considered to represent extreme wet (dry) events.

STANDARDIZING THE CALCULATIONS

Standardization of the procedure for computing the SPI is necessary so that all users will calculate index values which are comparable both spatially and temporally. If the same observed precipitation time series leads to different SPIs that depend on the computational procedures, then comparisons will not be of like quantities, and the comparisons will be confusing or misleading.

Recognizing that many users of the SPI may want to compute the index themselves, an intent of the standardization is to provide a "black box" software package for which the input is a precipitation time series and for which the output is the SPI. Ease of computation by a user is therefore a primary goal. Since many users may not have either the background, understanding or time to make statistical or climatological decisions regarding which probability density functions best describe the input data, a second goal of the standardization is to select one probability distribution function for all sites and time scales.

These two goals can be reached by answering the following questions:

1. How do SPI values change when different probability distribution functions are used?
2. Does a description of an event depend on the probability distribution functions?

The first question is concerned more with "scientific" considerations than with applications of the SPI, whereas the second question is concerned more with the applications. The answers to these questions should result in a standard procedure for computing the index.

DISTRIBUTION COMPARISONS

The precipitation data used in this study are byproducts of the development of the National Electronic Drought Atlas (Teqnical Services, 1997). They are also the same data that were used in a comparison of the SPI with the PMDI. This comparison resulted in a recommendation that the SPI be used as a primary drought index (Guttman, 1997). Observed monthly precipitation time series for 1,035 sites spread throughout the contiguous U.S. were the basis for computing time series of 1-, 3-, 6-, 12-, 24- and 36-month running precipitation totals. The record lengths vary among the sites, but were at least 60 years at all sites and averaged 85 years. Figure 1 is a histogram of the record lengths.

The Atlas contains probabilities of monthly, multimonthly, annual and multiyearly precipitation amounts. The probability densities were based on a regional, index flood procedure (Kite, 1988) and L-moment analyses. This resulted in regional probability distributions for 111 precipitation regions, 108 of which are homogeneous with respect to precipitation climate and frequency distribution characteristics (Guttman, 1993; Guttman *et al.*, 1993). The probability distributions that define the regional growth curves (cumulative probability curves) are not the same for all regions and are based on a best-fit analysis. For each of the 1,035 sites, the regional growth curves were scaled, the probabilities of the observed running precipitation totals were calculated, and then these probabilities were transformed into an SPI value.

The probabilities for a given site that were calculated from the above procedures (L-moment, regional analysis approach) include information from sites other than from the given site. It is assumed that since the collection of information comes from sites with similar characteristics (i.e., the sites are homogeneous), then the probability estimates for the given site are likely to be more accurate than estimates resulting from a single-site analysis. Therefore, the probabilities of observed precipitation for a given site that are calculated from the regional analyses are considered to be the best estimates of the true, underlying frequency distributions at the given site. A more complete description of the regional analysis procedure and of L-moments is given by Hosking and Wallis (1997).

SPI values for a site that are calculated from the regional probability analyses would be the most accurate, but the calculation requires an assessment of regional precipitation climate homogeneity so that the appropriate regional probabilities can be used. This assessment is difficult and time consuming, and is beyond the scope of the "ease of use" goal in the standardization of the SPI computations. The regional (REG) analysis approach does, however, result in the best SPI values to which values calculated with other methods can be compared. This study reports the

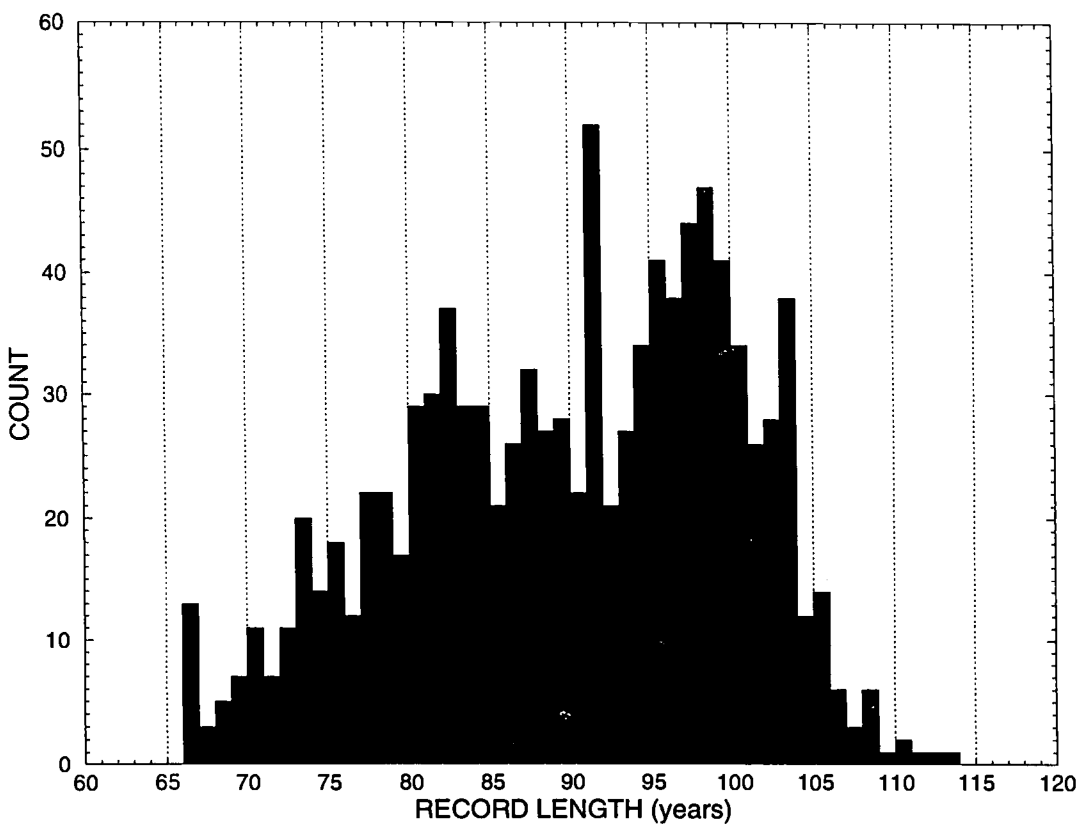

Figure 1. Histogram of Record Lengths.

comparison of SPI values for a site computed from a single station analysis and a common probability model to the SPI values for the site computed from the regional analysis.

Candidate distributions that are compared are the 2-parameter gamma (GAM), the 3-parameter Pearson Type III (PE3), which is sometimes referred to as the 3-parameter gamma, the 3-parameter generalized extreme value (GEV), the 4-parameter Kappa (KAP) and the 5-parameter Wakeby (WAK). For each of the 1,035 sites, data were stratified by calendar month, and L-moments were computed from the time series of 1-, 3-, 6-, 12-, 24- and 36-month precipitation totals. In the computation of the L-moments, observations of annual and shorter time scale total precipitation were assumed to be independent, as were non-overlapping multiyear totals (e.g., 24-month precipitation accumulated from months 1-24, 25-48, etc.). For each site and time scale, the L-moments were used to calculate the parameters of the above distributions. In addition, the gamma distribution used by McKee *et al.* (T. B. McKee, N. J. Doeskin, and J. Kleist, 1997, personal communication) was compared. This distribution is the 2-parameter gamma (CSU) for which the parameters were estimated by the maximum likelihood method. For the parameter estimations, the data were stratified by calendar month, and samples include overlapping multiyear totals (e.g., 24-month precipitation accumulated from months 1-24, 12-36, 24-48, etc.). The CSU model is currently used in the spatial portrayal of the SPI that is promulgated by, for example, the National Drought Mitigation Center (NDMC).

The comparison methodology was evolutionary. The first step was to visually inspect scatter diagrams of SPI values computed from the REG vs. candidate distribution models separately for each time scale. The effort was exploratory and was an attempt to gain an appreciation for the patterns displayed. Since each scatter diagram contained on the order of a million points, discerning patterns was difficult. Because many users of the SPI will be looking at color-coded spatial displays similar to the maps accessible through the NDMC internet website, it was decided to plot scatter diagrams only for REG and candidate distribution model values that were in different classes relating to the severity of an event. The SPI classes follow McKee *et al.* (1995) and are shown in Table 1. Considering only instances when the classes were different for the two models resulted in about a 90 percent reduction in the number of points that were plotted.

The scatter diagram for the GEV vs. REG model one-month SPI values is shown as an example in Figure 2. It is clear from this figure that patterns for wet (SPI > 0) events may not be the same as for dry (SPI < 0) events, and also that the pattern of positive differences (REG – candidate distribution) may not be same as the pattern of negative differences. Subjective evaluation of the scatter diagrams led to four characteristics for which the models would be assessed: (1) symmetry of differences between the regional and candidate distribution model values, (2) the magnitude of the differences, (3) the temporal variability of the differences, and (4) the spatial variability of the differences. A candidate distribution with the fewest number of differences, with the most symmetrical pattern of differences, with the smallest magnitude of differences, and with spatially and temporally invariant differences would be considered to be the "best" of the candidate distributions.

The evaluations are somewhat subjective because there is no perfect standard by which to compare the candidate distributions. While the REG model is considered to be the most correct, it is known a priori that there are some problems with regional homogeneity, distribution fitting, etc. (Guttman, 1993). This knowledge is subjectively incorporated into the current analysis. The imperfect standard also led to a simple comparison methodology: the characteristics for each of the candidate distributions were ranked in order of desirability from most to least desirable.

When model values were not in the same class (Table 1), differences between the REG and candidate distribution index values were examined for all time scales and for all sites. The number of differences, the mean difference, and the standard deviation of the differences were tabulated for all differences, for only wet events, and for only dry events. The three variables are indicators of the symmetry characteristic. The results were then ranked from 1 to 6 with 1 being the most desirable, and the ranks were summed. Summing the ranks implies that all three variables are equally important in determining the "best" model. Table 2 shows the results of the rankings. For all the differences, the PE3 and WAK models are considered to be better than the other candidate models, for wet events the GAM and PE3 models are better, and for dry events the PE3 and KAP models are better. The table also indicates when a characteristic for a candidate model is much worse (subjectively determined) than the other candidates.

The number of differences, N, is the ratio of the number of positive differences divided by the total number of differences. If N is equal to one-half, then the number of differences are considered symmetric. The absolute value of the departures of N from one-half, for the six distributions, for all events, ranges from .01 to .14 for the 1-month SPI. For longer durations, the departures range from .00 to .12 for the 3-month SPI down to .00 to .07 for the 36-month SPI. For wet events, the departures range from .00 to .16

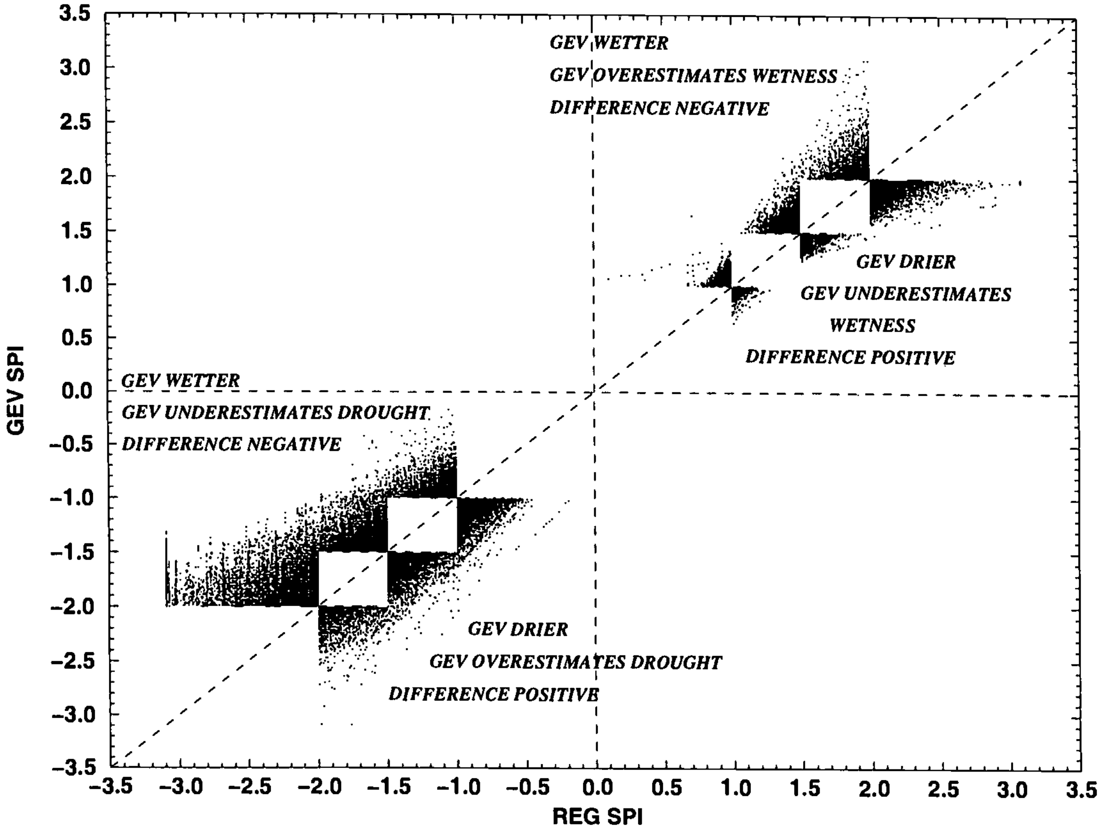

Figure 2. Scatter Diagram of GEV vs. REG Model One-Month SPI Values When Simultaneous Model Values Are in Different Event Classes.

for the 1-month SPI, and from .00 to .07 for the longer duration SPIs. For dry events, the departures range from .02 to .15 for all the SPIs.

TABLE 1. SPI Classes.

		SPI	≤	-2.0	Extremely Dry
-2.0	<	SPI	≤	-1.5	Moderately Dry
-1.5	<	SPI	≤	-1.0	Dry
-1.0	<	SPI	<	1.0	Neutral
1.0	≤	SPI	<	1.5	Wet
1.5	≤	SPI	<	2.0	Moderately Wet
2.0	≤	SPI			Extremely Wet

The absolute value of the differences in the means (MN) for all events ranges from .01 to .08 for the 1-month SPI. For longer durations, the range is from zero to .04. For wet events, the range is from .00 to .07 for shorter duration SPIs, from .00 to .10 for the 24-month SPI, and from .01 to .15 for the 36-month SPI.

The absolute value of the differences in the standard deviations (SD) for all events range from zero to .11 for the 1-month SPI down to zero to .07 for the 36-month SPI. For wet events, the range varies from .00 to .12 for the shorter duration SPIs, from .02 to .15 for the 24-month SPI, and from .05 to .20 for the 36-month SPI. For dry events, the ranges are from .00 to .15 for the 1- and 3-month SPIs, .02 to .08 for the 6- and 12-month SPIs, .00 to .10 for the 24-month SPI, and .00 to .16 for the 36-month SPI.

The REG – candidate distribution differences when index classes are not the same were further stratified by the sign of the differences. Positive (negative) differences indicate that the REG model produces less (more) extreme SPI values during dry (wet) conditions than does a candidate distribution. Rankings of mean differences, variability of the differences, and the total number of differences (fewest is "best") are shown in Table 3. The five variables in the table are indicators of the magnitude characteristic.

The range of the values of these five variables over the distributions, stratified by duration and type of event, is given in Table 4. For all events, the total number of differences is generally higher for 1-month

TABLE 2. Distribution Comparison Rankings for Number of Differences (N), Mean Differences (MN) and Standard Deviation of Differences (SD) for All, Wet and Dry Event Differences (REG – candidate distribution SPI value) When Index Classes Are Not the Same for REG and Candidate Distribution. Starred ranks indicate much worse than other distributions.

Event	Distribution	N	MN	SD	Sum of Ranks	Rank of Ranks
ALL	CSU	6*	5	2.5	13.5	5
	GAM	5	3	2.5	10.5	4
	GEV	4	6	6*	16	6
	KAP	2	4	5	11	3
	PE3	1	1	4	6	1.5
	WAK	3	2	1	6	1.5
WET	CSU	6	2.5	4	13.5	5
	GAM	2	1	2	5	1
	GEV	3	4	3	10	3
	KAP	1	5	5*	11	4
	PE3	4	2.5	1	7.5	2
	WAK	5	6	6*	17	6
DRY	CSU	6*	3	6	15	6
	GAM	5*	2	3	10	3
	GEV	3	5.5	4	12.5	4
	KAP	2	4	1	7	2
	PE3	1	1	2	4	1
	WAK	4	5.5	5	14.5	5

TABLE 3. Distribution Comparison Rankings for Number of Positive vs. Negative Differences (N), Mean Positive (MN+) and Negative (MN-) Differences and Standard Deviation of Positive (SD+) and Negative (SD-) Differences for All, Wet and Dry Event Differences (REG – candidate distribution SPI value) When Index Classes Are Not the Same for REG and Candidate Distribution. Starred ranks indicate much worse than other distributions.

Event	Distribution	N	MN+	MN-	SD+	SD-	Sum of Ranks	Rank of Sums
ALL	CSU	4	3	1.5	1	1	10.5	2
	GAM	1	2	1.5	2	2	8.5	1
	GEV	2	1	5	3	5.5	16.5	3
	KAP	5	4.5	4	5	5.5	24	5
	PE3	3	4.5	3	4	3	17.5	4
	WAK	6*	6*	6	6*	4	28	6
WET	CSU	3	6*	2	6	2	19	3
	GAM	1	1.5	1	3	1	7.5	1
	GEV	4	3	4	5	4	20	4
	KAP	5	4	5*	2	6*	22	5
	PE3	2	1.5	3	4	3	13.5	2
	WAK	6*	5	6*	1	5*	23	6
DRY	CSU	3	1	1.5	1	2	8.5	1
	GAM	1	3	1.5	3	1	9.5	2
	GEV	2	2	3	2	5	14	3
	KAP	5	4	4	5*	3.5	21.5	4
	PE3	4	5	6	4*	6	25	5
	WAK	6*	6*	5	6*	3.5	26.5	6

than 3-, 6-, or 12-month durations, but still less than for the 24- and 36-month durations. For each event, the mean positive differences and standard deviations generally have the same ranges for the 1- through 12-month durations with higher values for the longer durations. This same pattern is also evident for the negative differences and standard deviations, except that when stratified by event, the shorter durations exhibit slightly higher values than the moderate durations.

TABLE 4. Range of Values Over the Distributions for Characteristics in Table 3, Stratified by Duration (months).

Duration	N (thousands)			MN+			MN-			SD+			SD-		
	All	Wet	Dry	All	Wet	Dry	All	Wet	Dry	All	Wet	Dry	All	Wet	Dry
1	102-106	35-41	63-69	.17-.22	.14-.18	.20-.29	.17-.22	.20-.27	.22-.34	.13-.24	.11-.17	.12-.27	.13-.24	.13-.21	.18-.29
3	92-104	37-43	55-61	.17-.22	.14-.18	.20-.29	.17-.22	.14-.23	.21-.28	.13-.24	.11-.17	.12-.27	.13-.24	.13-.21	.18-.29
6	89-103	38-45	47-58	.17-.22	.14-.18	.20-.29	.17-.22	.14-.23	.18-.23	.13-.24	.11-.17	.12-.27	.13-.24	.09-.20	.13-.26
12	87-105	42-47	46-57	.17-.22	.14-.18	.20-.29	.17-.22	.14-.23	.18-.23	.13-.24	.11-.17	.12-.27	.13-.24	.09-.20	.19-.26
24	113-145	53-67	60-78	.22-.33	.20-.23	.23-.41	.22-.33	.19-.33	.25-.29	.15-.34	.11-.17	.13-.35	.15-.34	.12-.36	.19-.26
36	131-171	63-80	68-91	.26-.41	.25-.28	.26-.50	.26-.41	.23-.42	.28-.33	.15-.34	.21-.23	.15-.41	.15-.34	.14-.41	.19-.26

The rankings provide quantitative comparison information. A qualitative assessment, however, of all the scatter diagrams, counts, mean differences and variability of the differences led to several general observations that encompass the characteristics. For all candidate distributions, first, the mean differences between the index values and the variability of these differences are between one and two times greater for dry events than for wet events. Second, the means and variability are greater for 24- and 36-month time scales than for 1- to 12-month time scales. Third, the number of differences are greater for the 1-, 24- and 36-month time scales than for the 3- to 12-month time scales. The larger means, greater variability, and greater number of differences for the multiyear time scales results from estimating probability densities from small sample sizes. The number of independent 24- and 36-month samples are one-half and one-third, respectively, that of annual samples. Guttman (1994) investigated the effect of sample size on the central tendency, variability, skewness and kurtosis of distributions of non-zero precipitation fit by L-moments and concluded that sample sizes of least 60 should be used to estimate the tails of a distribution; the smaller the sample size, the more unreliable the probabilities become. The greater number of differences for the 1-month time scale results from the much higher variability of monthly precipitation amounts than for the longer time scales.

It was also noted that the pattern of differences is most symmetrical for the PE3 distribution and least symmetrical for the GEV, and that the magnitude of the differences is smallest for the GAM distribution. The KAP and WAK models tended to fit the site data well, but often did not agree with the REG model. This observation is not surprising since a 4- or 5-parameter model is generally expected to fit the observed data well. However, because of the sampling variability and homogeneity assumptions implicit in the regional model, it may not adequately reflect the regional growth curve.

Spatial and temporal comparisons were made by assessing histograms and scatter diagrams of the index value differences when the class of the candidate distribution was not the same as that of the REG model. The data were stratified both by dry and wet events and by positive and negative differences, as well as by month and region. The histograms were of the frequency of sites by month with at least one difference. The scatter diagrams were of combinations of the number of differences by site, the average difference by site, month, and region.

The histograms show a uniformity of frequencies throughout the year for all distributions, time scales, and stratification. Scatter plots of the average site difference vs. month and the number of differences by site vs. month also show a uniformity throughout the year. The one exception to this uniform temporal pattern is a high number of differences for months at sites where precipitation falls infrequently.

This exception is depicted in Figure 3, which shows a spatial pattern of the number of differences per site. Note the cluster of points with frequencies above 60. All of these points are for desert locations and reflect counts in months for which the number of non-zero precipitation amounts during the period of record is less than 17. Other than this exception, the scatter diagram is typical of the spatial uniformity exhibited by all the other similar scatter diagrams. In addition to the number of differences, the average site differences are also spatially invariant.

Qualitatively, the general histogram and scatter diagram patterns do not indicate any dependence on time of year. Also, with the exception just noted, they do not show any differences among geographical regions.

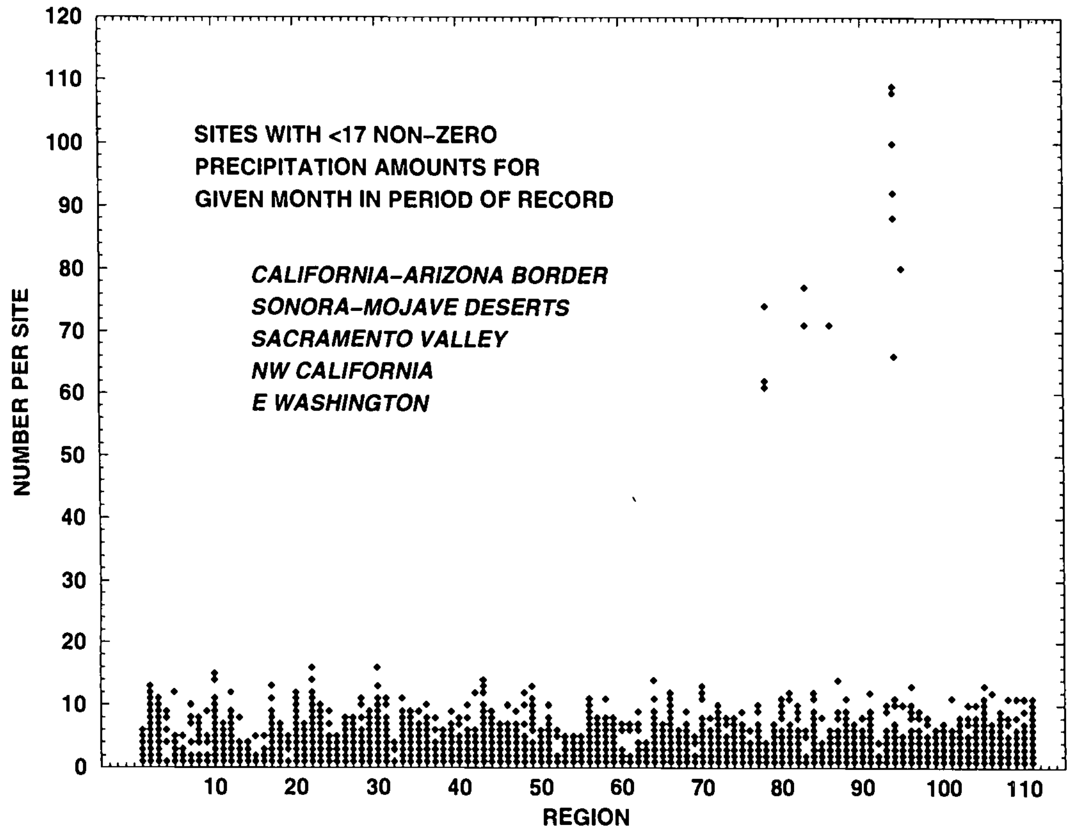

Figure 3. Scatter Diagram of the Number of Differences Per Site, by Region, Between One-Month, Wet Event SPI Values Computed from PE3 and REG Models When Simultaneous Model Values Are in Different Classes and When the PE3 SPI Value is Higher Than the REG SPI Value.

DISTRIBUTIONS AND DRY EVENTS

Realistically, the effect a specified distribution has on the number, length and intensity of dry events is perhaps more important to users of the SPI than the distribution comparison described in the previous section. Although the SPI also measures wet events equally effectively as dry events, most long-lasting anomalously high precipitation events are predicted well in advance, whereas dry events are more insidious and difficult to predict. Therefore, this section does not consider the effects different distributions have on wet event descriptions.

Dry events are defined as the sequence of months in which the SPI is continuously negative with the condition that the SPI is less than or equal to -1 in at least one of the months in the sequence. For each site for each distribution and time scale, the number of dry events was tabulated. Because the period of record of the precipitation data varies among the sites, the number was transformed into the standardizing measure of an event rate that is equal to the number of events per 100 years. The length in months of each event as well as the intensity of each event were tabulated. The intensity is defined as the absolute value of the sum of the monthly SPI values from the beginning to the end of an event and has units of months (McKee *et al.*, 1993).

The mean and standard deviation of the differences of the event rate, length and intensity between each candidate distribution calculation and the REG model calculations were computed and ranked similarly to the distribution comparison methodology. The results are shown in Table 5.

Scatter diagrams were plotted of the number of droughts per 100 years resulting from candidate distribution calculations vs. the number resulting from the REG model. Figure 4 is an example and is typical of all the distribution and time scale diagrams. The diagonal line in the figure represents perfect correspondence between the models; the scatter of points indicates an excellent relationship. Histograms of the differences in event rates (e.g., Figure 5) confirm the close agreement between the candidate and REG model rates. Spatially, Figures 6 and 7 are examples which show the few locations where the differences are greatest. The scale for these figures indicates the differences in terms of the number of years in which one more or one fewer event can be expected to occur.

TABLE 5. Dry Event Comparison Rankings for Mean Differences (MN) and Standard Deviation of Differences (SD).

Characteristic	Distribution	MN	SD	Sum of Ranks	Rank of Sums
RATE	CSU	6	4	10	6
	GAM	4	1	5	2
	GEV	5	3	8	4
	KAP	3	5	8	4
	PE3	1	2	3	1
	WAK	2	6	8	4
LENGTH	CSU	5	4.5	9.5	6
	GAM	6	1	7	3.5
	GEV	2	2.5	4.5	1
	KAP	3.5	4.5	8	5
	PE3	3.5	2.5	6	2
	WAK	1	6	7	3.5
INTENSITY	CSU	2.5	6	8.5	4.5
	GAM	4	1	5	1.5
	GEV	6	2.5	8.5	4.5
	KAP	5	4	9	6
	PE3	2.5	2.5	5	1.5
	WAK	1	5	6	3

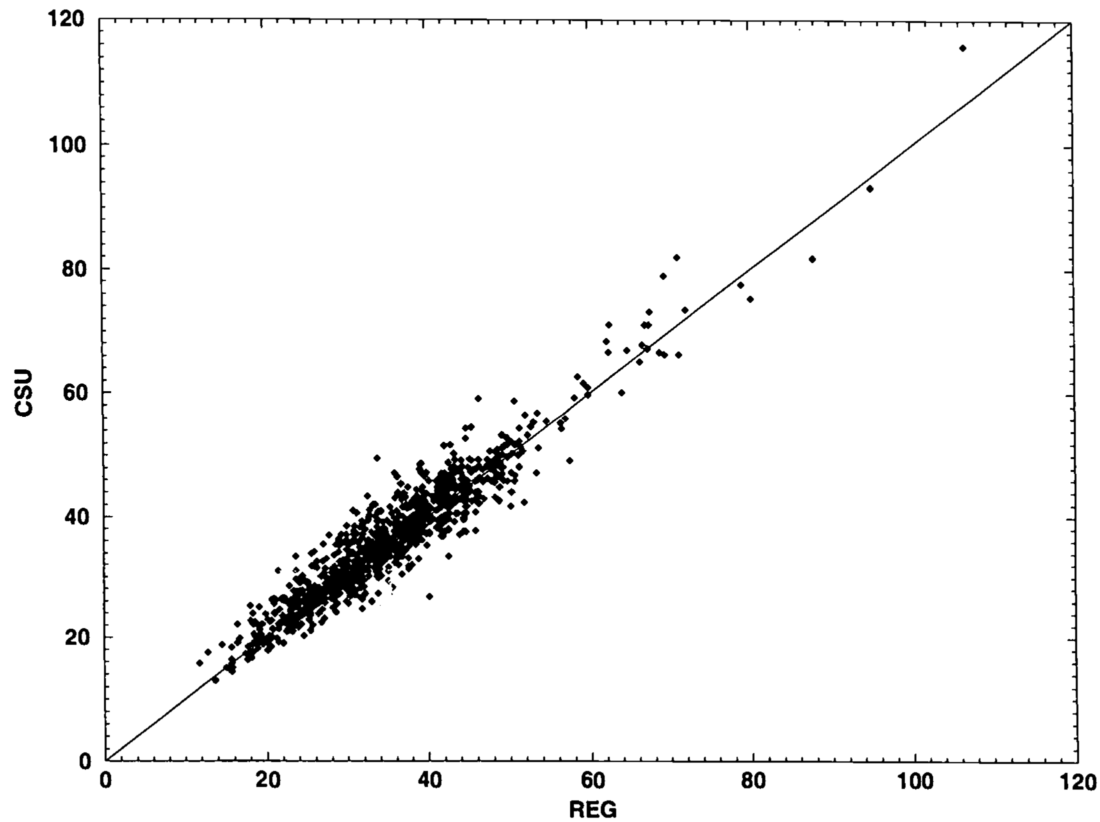

Figure 4. Scatter Diagram of the Number of Droughts Per 100 Years Computed from the CSU and REG Six-Month SPI Models. Diagonal line indicates perfect agreement between the two models.

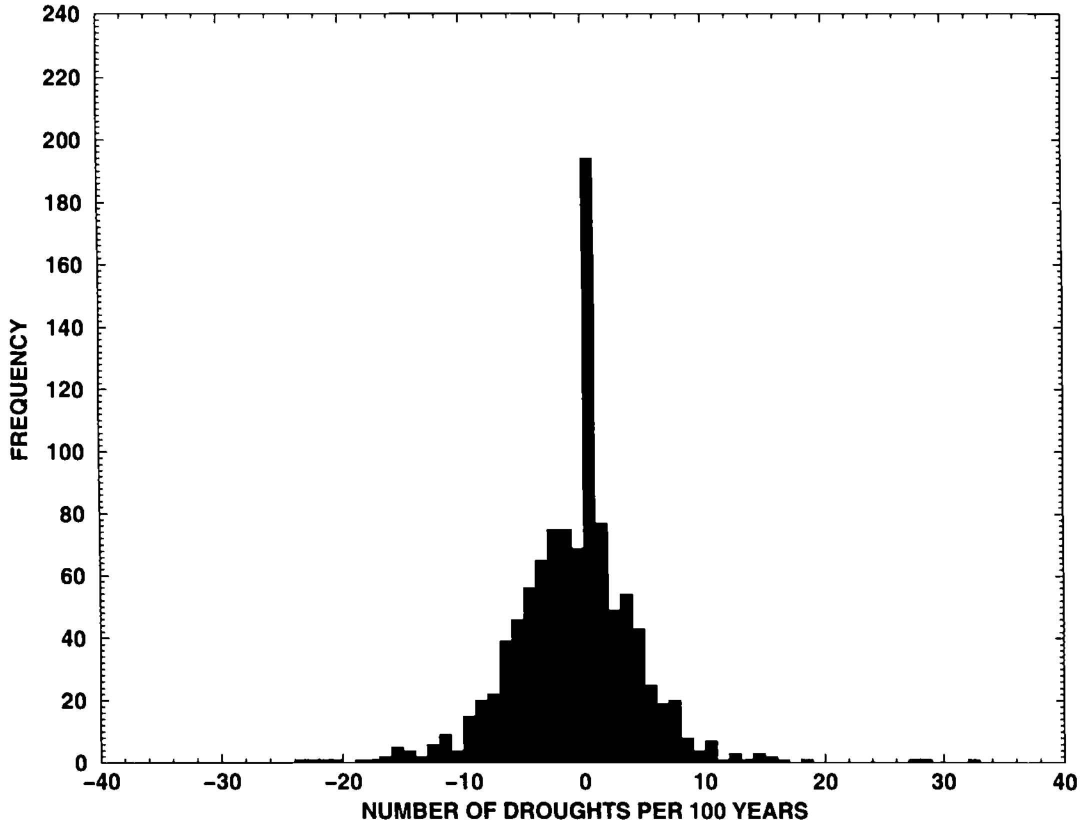

Figure 5. Histogram of the Differences in Dry Event Rates Computed from PE3 and REG Three-Month SPI Values.

After looking at all the histograms, scatter diagrams and tabular information regarding the number of droughts, for almost all candidate distributions and for all time scales, the event rate is greater than that resulting from the REG model. The magnitude of the difference, however, is small. For the 1-month SPI, there are about 2 more dry events per 100 years from the candidate distributions than from the REG model, and for longer time scales, there is about one more event every few hundred years. There is not any readily apparent spatial pattern to the results.

Graphs of the histograms, scatter diagrams and tabular information regarding the length and intensity of dry events were similarly assessed. For almost all candidate distributions and time scales, the mean event length is slightly greater for the REG model calculations. The mean difference is about .1 month for events measured by SPI values with time scales of a year or less, and between a .5 and 1 month for events measured by SPI values with 2- and 3-year time scales. The mean intensity as measured from the candidate distribution calculations is nearly the same as that from REG model, with differences being less than .1 month for time scales of a year or less and about .5 month for the longer time scales. Spatially, the results are consistent at all sites.

DISCUSSION AND CONCLUSIONS

For dry events, the assessment shows that there is very little difference in the number, duration, intensity or regional variation of the events resulting from the computations using the compared distributions. Based on this assessment alone, it does not matter which distribution is chosen. However, the assessment of distribution comparisons shows that the PE3 and GAM models perform the best in comparison to the "standard" REG model. Because the PE3 with its three parameters is more flexible than the GAM with its two parameters in fitting sample data, it is suggested that the PE3 be used as a universal model from which to compute the SPI from site data. Although not discussed, the PE3 model is also the best choice for describing wet events.

The distribution analysis provides some guidelines for determining how many data are needed to adequately estimate the probabilities. The mean departures from the standard REG model, as well as the variability of the departures, is greater for all compared distributions for the 24- and 36-month time scales than for the shorter time scales. Since the sample size of the 24- and 36-month time scales is a half and a third, respectively, of that for the 1- to 12-

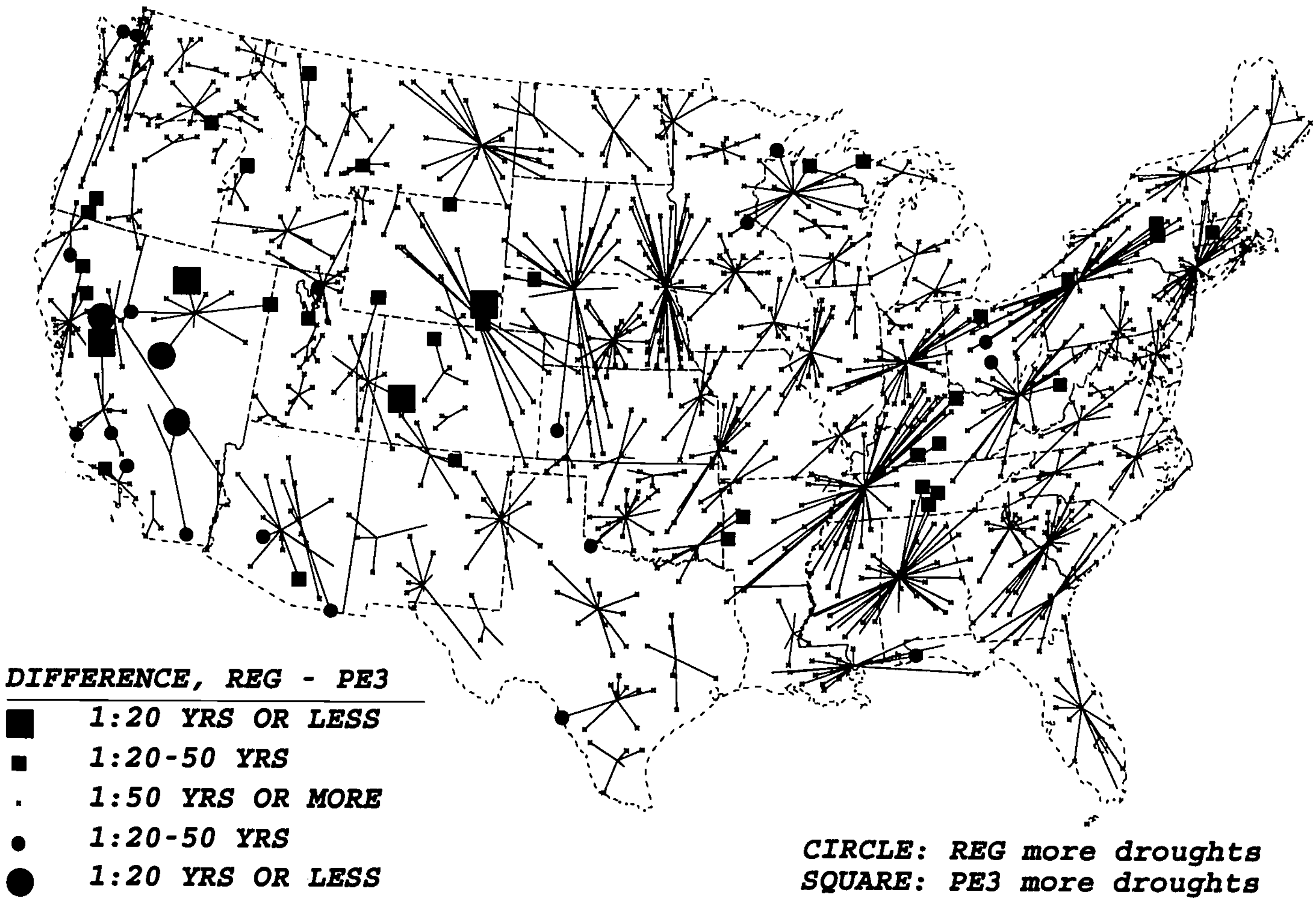

Figure 6. Difference in the Number Per 100 Years of Three-Month SPI Dry Events Computed from PE3 and REG Models. Lines radiate from the center of a region to the location of a site.

month time scales, the reliability of the probability estimates can be considered to be function of sample size. Further supporting this contention is the cluster of outliers shown in Figure 3. All of the points in the cluster are based on probabilities of non-zero precipitation amounts derived from very small samples.

The period of record for most sites is of the order of 80-100 years, so that the sample size for computing multiyear SPI values is 40-50 or less. These sample sizes are smaller than those recommended by Guttman (1994). Operationally, however, extreme events such as 3-year or longer droughts are of major importance to decision makers. It is therefore suggested that as a compromise between the scientific issues and the practical use of a drought index, the SPI be computed for time scales of 1 to 24 months so reasonable probability estimates can be obtained, and so the events of longer duration can be monitored by analyzing the time series of both overlapping and non-overlapping 24-month SPI values.

Selecting the PE3 distribution as a universal model for calculating the SPI is not enough for the SPI to replace the Palmer indices as the primary drought index. Algorithms that calculate probabilities from the PE3, as well algorithms that calculate inverse normal functions, vary. It is therefore necessary to provide a user friendly, standard software package for which the computational complexities are transparent to the user. Such a package is now available electronically via file transfer protocol (ftp) from an internet browser at URL:

ftp://www.ncdc.noaa.gov/pub/software/palmer/spinew.f

This file is a Fortran 77 source program that computes the SPI for time scales ranging from 1 to 24 months. The PE3, using L-moment algorithms, is used as the probability model. There are several output options available to the user. Copious remarks are embedded in the program, and an internet URL is given in the program for accessing monthly precipitation data.

The tools are now available for operationally calculating the SPI. It is strongly suggested that users of

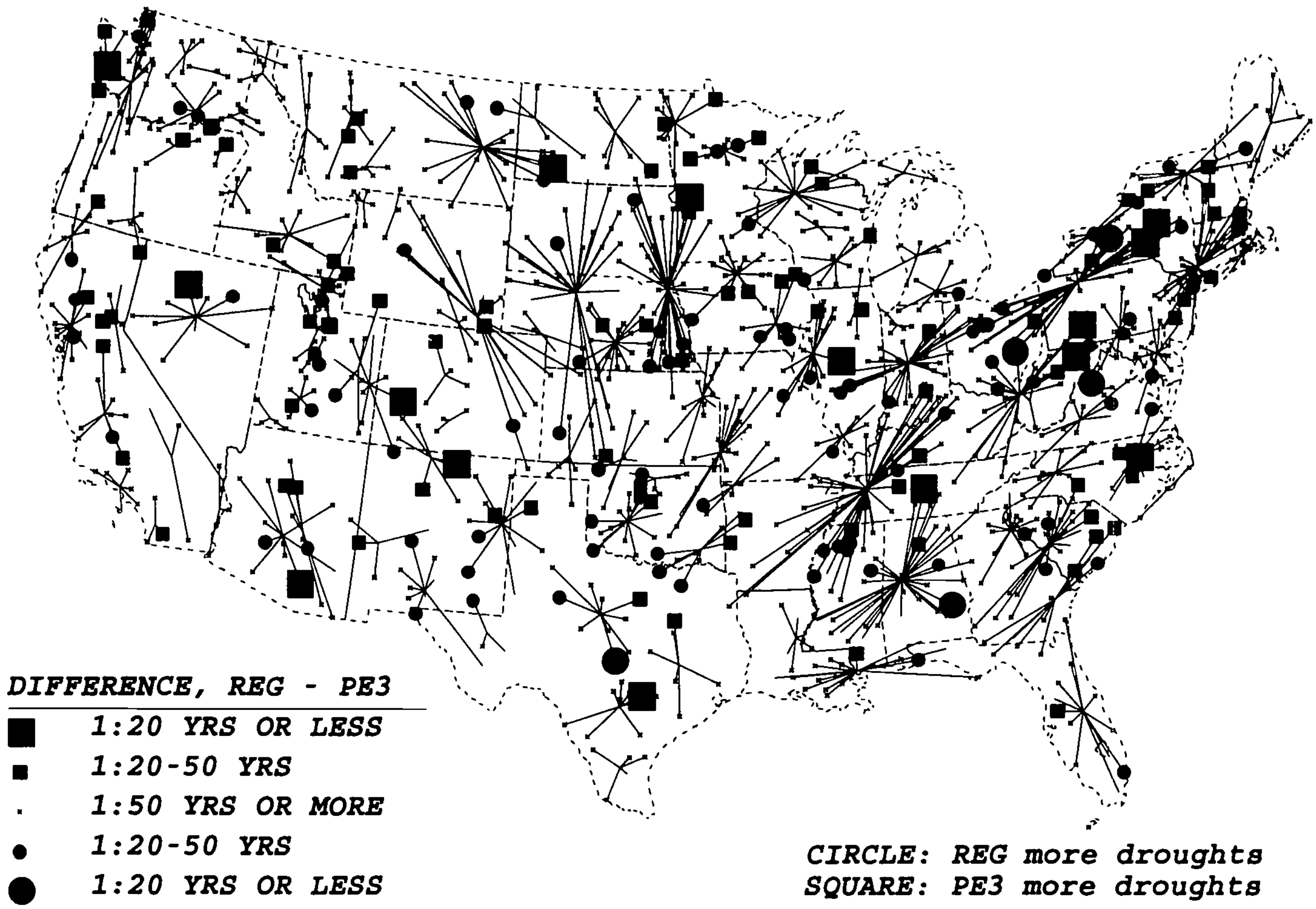

Figure 7. Difference in the Number Per 100 Years of Twelve-Month SPI Dry Events Computed from PE3 and REG Models. Lines radiate from the center of a region to the location of a site.

the Palmer indices consider using the SPI as either a primary drought index or as an equal companion to the Palmer indices.

LITERATURE CITED

Alley, W. M., 1984. The Palmer Drought Severity Index: Limitations and Assumptions. J. Clim. Appl. Meteor. 23:1100-1109.

Guttman, N. B., 1991. Sensitivity of the Palmer Hydrologic Drought Index. Water Resources Bull. 27:797-807.

Guttman, N. B., 1993. The Use of L-Moments in the Determination of Regional Precipitation Climates. J. Clim. 6:2309-2325.

Guttman, N. B., 1994. On the Sensitivity of Sample L Moments to Sample Size. J. Clim. 7:1026-1029.

Guttman, N. B., 1997. Comparing the Palmer Drought Index and the Standardized Precipitation Index. J. Amer. Water Resources Assn. 34:113-121.

Guttman, N. B., J. R. M. Hosking, and J. R. Wallis, 1993. Regional Precipitation Quantile Values for the Continental U.S. Computed from L-Moments. J. Clim. 6:2326-2340.

Guttman, N. B., J. R. Wallis, and J. R. M. Hosking, 1992. Spatial Comparability of the Palmer Drought Severity Index. Water Resources Bulletin 28:1111-1119.

Heddinghaus, T. R. and P. Sabol, 1991. A Review of the Palmer Drought Severity Index and Where Do We Go From Here? Proc. 7th Conf. on Applied Climatology, September 10-13, 1991, American Meteorological Society, Boston, Massachusetts, pp. 242-246.

Hosking, J. R. M. and J. R. Wallis, 1997. Regional Frequency Analysis. Cambridge Univ. Press, 224 pp.

Karl, T. R., 1983. Some Spatial Characteristics of Drought Duration in the United States. J. Clim. Appl. Meteor. 22:1356-1366.

Karl, T. R., 1986. The Sensitivity of the Palmer Drought Severity Index and Palmer's Z-Index to Their Calibration Coefficients Including Potential Evapotranspiration. J. Clim. Appl. Meteor. 25:77-86.

Kite, G. W., 1988. Frequency and Risk Analyses in Hydrology. Water Resources Publications, Littleton, Colorado, 257 pp.

McKee, T. B, N. J. Doeskin, and J. Kleist, 1993. The Relationship of Drought Frequency and Duration to Time Scales. Proc. 8th Conf. on Applied Climatology, January 17-22, 1993, American Meteorological Society, Boston, Massachusetts, pp. 179-184.

McKee, T. B, N. J. Doeskin, and J. Kleist, 1995. Drought Monitoring with Multiple Time Scales. Proc. 9th Conf. on Applied Climatology, January 15-20, 1995, American Meteorological Society, Boston, Massachusetts, pp. 233-236.

Palmer, W. C., 1965. Meteorological Drought. Res. Paper No. 45, Weather Bureau, Washington, D.C., 58 pp.

Teqnical Services, 1997: National Electronic Drought Atlas. CDROM, New London, Connecticut.